DOCUMENT RESUME

ED 071 626 HE 003 782

AUTHOR	Cope, Robert G.; And Others
TITLE	A Study of the Relationship Between Quality Instruction as Perceived by Students and Research Productivity in Academic Departments. Final Report.
INSTITUTION	Washington Univ., Seattle. Coll. of Education.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C. Regional Research Program.
BUREAU NO	BR-1-J-010
PUB DATE	Nov 72
CONTRACT	OEC-X-72-0021
NOTE	45p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Educational Quality; Evaluation Methods; *Higher Education; Institutional Research; *Research; *Teacher Evaluation; *Teaching Quality

ABSTRACT

This report presents the findings of a study of the relationship between quality teaching and reputation for research and publication in 17 academic departments at the University of Washington. A specially prepared student rating of teaching form was administered in the spring of 1972. The results from the student ratings of instruction were compared with these departments' national ranking in the most recent American Council on Education rating of graduate programs. The results did not indicate that a relationship exists between the student perception of quality of instruction and the reputation departments have among peers for research and publications. The results did show, however, that the relationship between student ratings and research reputation differed according to field of study, and that the results of current research do enter into the instruction received in advanced courses. (Author)

ED 071626

1-J-010
NCERT
HE

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FINAL REPORT

A Study of the Relationship Between Quality Instruction as Perceived by Students and Research Productivity in Academic Departments

Project No. 1–J–010 Grant No. OEC–X–72–0021

Robert G. Cope, John G. McMillin, Judy M. Richardson
University of Washington, College of Education, Seattle, WA 98195

November 1972

U.S. Department of Health, Education, and Welfare Office of Education
National Center for Educational Research and Development (Regional Research Program)

HE003782

ABSTRACT

This report presents the findings from a study of the relationship between quality teaching and reputation for research and publication in seventeen academic departments at the University of Washington. A specially prepared student rating of teaching form was administered in the spring of 1972. The results from the student ratings of instruction were compared with these departments' national ranking in the most recent American Council on Education rating of graduate programs.

The results did not indicate a relationship exists between the student perception of quality of instruction and the reputation departments have among peers for research and publication. The results did show, however, that the relationship between student ratings and research reputation differed according to field of study, and that the results of current research do enter into the instruction received in advanced courses.

FINAL REPORT

Project No. 1-J-010 Grant No. OEC-X-72-0021

A Study of the Relationship Between Quality Instruction as Perceived by Students and Research Productivity in Academic Departments

Robert G. Cope, John G. McMillin and Judy M. Richardson

University of Washington
Seattle, Washington

November 1972

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education
National Center for Educational Research and Development
(Regional Research Program)

TABLE OF CONTENTS

Objective and Introduction 1

Description of the Study 4

- Quality of Graduate Faculties 5
- Measuring Quality in a Subjective Survey 10
- Respondents 10

Analysis 12

- The Scales: Internal Consistency 12
- Mean Scale Ratings by Level of Instruction 17
- Relationship Between A.C.E. Rank and Student Ratings 17

Summary 25

Implications 25

References 27

Appendices

- A Sample Page from Roose and Anderson, 1970
- B Eidsmoe Scale
- C Cover Letter
- D Scale for Student Assessment of Teaching
- E Staff Deployment by Department
- F Details on Scale Means

Acknowledgements

We circulated an earlier summary of this research to the chairmen in the 17 departments and found these colleagues to be conscientious and generous with their criticisms, and helpful in correcting misinterpretations; we are indebted to them for this. We have tried to take account of their comments, but as one might expect there are still points with which they will disagree, as we disagree among ourselves.

Keith and Alice Pailthorp, Jan McMillin, Claudette Cope, and Toshio Nishi helped with the mailing of the questionnaire. Tom Langen offered constructive suggestions. Terry Eade compiled the "demographic data" on the departments, and Wilma Boyd typed the final report. We express our appreciation to them.

R.C. J.M. J.R.
Seattle, November 1972

LIST OF TABLES

Table

1.	Rank of University of Washington Departments Based Upon A.C.E. Ratings (1970)	6
2.	Rank of University Departments by Area of Knowledge	7
3.	Response Rates	11
4.	Correlation Matrix All Cases	13
5.	Correlation Matrix Lower Division	14
6.	Correlation Matrix Upper Division	15
7.	Correlation Matrix Graduate	16
8.	Mean Scale Ratings for Instruction in Lower, Upper, and Graduate Level Courses	18
9.	Mean Student Ratings for Departments	20
10.	Mean Student Ratings for Social Science Departments	22
11.	Rank Order of Departments by A.C.E. Rank and Department Means	23
12.	Rank Order of Departments by Areas of Knowledge, A.C.E. Rankings and Mean Student Ratings	24

LIST OF FIGURES

Figure

A. Scale for Student Assessment of Teaching 9

B. Cluster Analysis of Data from Table 4 13

C. Cluster Analysis of Data from Table 5 14

D. Cluster Analysis of Data from Table 6 15

E. Cluster Analysis of Data from Table 7 16

F. Comparative Scale: Means by Levels of Instruction 18

G. Graph of Scale Means 22

A STUDY OF THE RELATIONSHIP BETWEEN QUALITY INSTRUCTION AS PERCEIVED BY STUDENTS AND RESEARCH PRODUCTIVITY IN ACADEMIC DEPARTMENTS

Objective

The purpose of this study is to examine the proposition that the quality of instruction is related to the quality of research in academic departments at the University of Washington.

Introduction

The Relationship Between Research and Teaching: Opinions

Sanford (1967) and Jencks and Riesman (1968) have among others commented upon the relationship existing between the graduate school with its research orientation and the teaching function.[1] Sanford stated, "More and more, colleges have come to resemble graduate schools. As a research specialist, the college teacher has the same interests as his colleagues in graduate school and naturally seeks to make his students resemble graduate students as early as possible, to the neglect of their general development. In this situation, teaching becomes a lost art...." (Sanford, 1967, p. 169). Jencks and Riesman are more to the point, "What the graduate schools define as 'research' will get done; what they exclude is likely to languish" (p. 514).

The resulting national trend has been the downgrading of undergraduate instruction in favor of graduate education, the rush to the research project, increased consulting for business and government, and the published manuscript. Career conscious faculty members know that recognition, mobility, pay raises and promotions come with publication, and coversely there is little reward for quality in instruction. Desmond (1969) put it most succinctly, "The total impact of these problems is that the career interests of faculty are pitted squarely against the educational interests of students, especially the undergraduates"(p. 25).

Stern (1969) speaks directly to the matter of departmental curriculum and faculty affairs:

> Graduate departments require graduate faculties, and these tend to be a very select group even among the total population of faculty Ph.D.'s. They have been recruited for the past twenty years on the basis of their potential for research, publication, and program building. Having been screened even earlier for such qualities by their own graduate instructors, they are unquestionably the most aggressive, ambitious, energetic, counteractive,

---

[1] Cf. Sample, 1972; Anderson, 1968.

> pragmatic, and intellectually independent of all graduate school products, and committed both vocationally and by personal conviction to the development of others like themselves (p. 125).

Stern (1969) goes on to be less descriptive and more pointed in suggesting what the nature of the problem is in regard to the curriculum:

> What I am trying to suggest is that the graduate schools are a source of people, both students and faculty, who are: (1) independent of life outside the university community, and therefore more detached in their view of that world, (2) articulate and analytical, and therefore more likely to formulate a critical position on social issues, (3) engaged in a struggle paralleling that of the surgeons earlier in this century, for control of the institution that has become more and more specifically adapted (like the hospitals) to meet their particular professional needs, and (4) contributing inadvertently to a growing reservoir of frustration and ill-will among the enormously large numbers of students, graduate and undergraduate, who have neither the inclination nor the capacity to be included among the select few but who are nevertheless required to go through the same curriculum. It is after all not only the best curriculum, since it was designed to prepare people for graduate school, but also the only one (p. 126).

While the foregoing examples are drawn from comments made about higher education generally one need not look long before seeing evidence of the same concerns at the University of Washington. For example, in the spring of 1970, the Academic Vice President, writing for the student paper had this to say:

> The large university, the university which in recruiting and promoting its faculty very seldom inquires about success or potential success in teaching, but only about research productivity, is confronted by the allegation that its faculty is engaged in a mass 'flight from teaching.' Let us not deceive ourselves; as with similar institutions, the University of Washington is impaled on the uncomfortable horns of a dilemma. We profess to be first and foremost a teaching institution and indeed most of our support from the state government is primarily given to enable us to discharge this function, but much of our support from the federal government and from foundations, comes to us for research and without that research our reputation would be local and parochial.
>
> So today, as never before, scholarly research and publications flourish. Enticing grants and stipends from outside sources have sometimes enabled the professor, an airborne holder of multiple fiefs and benefices, to absent himself from teaching duties with a high degree of regularity to serve as a consultant or to attend workshops, colloquia, and conferences (p. 4).

Recognizing the importance of research, Soloman Katz, nevertheless, and consistent with current assumptions about the relationship of research and teaching, goes on to say:

> The dichotomy, the 'either . . . or' argument, which one finds vigorously presented in terms of the supposed mutual exclusiveness of teaching and research, I believe is false and dangerous. It turns on a false set of alternatives. It suggests that some faculty members publish while others are good teachers. But there is nothing incompatible between teaching and research. A member of the faculty may be heavily committed to research, enhancing his own and his institution's prestige. He need not thereby neglect his role as teacher in the revelation of knowledge from the introductory lecture course to the dialectic of the graduate seminar. The evidence of that devotion to teaching may not always be available, it may not always be as readily rewarded as his research, but unless he concerns himself with teaching, he will fall far short of fulfilling one of the major responsibilities of the calling to which he has dedicated himself (p. 4).

The Dean of the College of Arts and Sciences (presently Executive Vice President) in an address on the care of faculty entitled "The Economics of Deaning: The Care and Feeding of Homo Academicus," clearly indicated the basis on which faculty members are judged. Cartwright stated, "The criterion of quality of the faculty will be determined almost exclusively on the basis of the research scholarship of the faculty as evidenced through publication, exhibition, or performance (Cartwright, 1965, p. 53)."

Thus, while the possible conflict between research and teaching is recognized, the University, like nearly every institution of similar status, sees these as compatible functions, while placing the greatest emphasis on rewards for research and publication.

The Relationship Between Research and Teaching: Research

While the seemingly inherent conflict between research activities and the quality of teaching have been the subject of considerable discussion, empirical studies have been scarce. Dressel, et.al., in an investigation undertaken in 1967 of fifteen universities and their departmental[2] operation concluded:

> It was clear that many deans and chairmen felt the best way to produce quality undergraduate instruction was to develop a quality research and graduate program which would attract a good faculty. Yet universities that had the highest ratings on quality of graduate programs and the graduate faculty as judged by the Cartter report,[3] tended to place least emphasis on undergraduate instruction and showed lowest concern for students (Dressel, 1970, p. 43).

---

[2]Departments of mathematics, psychology, history, English, chemistry, business administration and engineering were selected for more intensive study within the universities.

[3]For a description of this study, also referred to as the ACE rating, see section "Quality of Graduate Faculties," p. 5 this study.

This and similar relationships were stated in a number of different ways throughout the book and are quoted here to emphasize the pervasiveness of the two orientations. For example Dressel, et.al., reported that "...the departments regarded as really good were invariably those that were able to demonstrate an active and current output of research publications, graduate students, receipt of federal monies, and receipt of fellowships from national organizations"(p. 22).

Dressel, et.al., also, not surprisingly, found that more publications per faculty member were positively related to higher ratings on the Cartter Report (1966)[4] but baccalaureates awarded were negatively related to the rate of publication per faculty number (p. 46).

It was also reported that "...faculty whose departments were mentioned favorably in the report [Cartter Report] placed a relatively lower emphasis on undergraduate instruction of both majors and minors. On the other hand, it is clear that departments mentioned in the report placed a relatively higher emphasis upon research, whether it be basic or applied"(underlining added, p. 50).

And finally, from Dressel, et.al., "Apparently undergraduate instruction and basic research represent two distinctive missions which go far to determine the entire character of a department. Some faculty perceived them as nearly antithetical; certainly high priority on one meant relatively low priority on the other"(p. 76).

Four other studies have focused more directly upon the central question. Bresler (1968), at Tufts University, in a study of teaching effectiveness comparing faculty who had acquired government research grants with those who had not, found that those who had acquired funds for research received higher teacher ratings. Hayes (1971) however, in a similar study at Carnegie-Mellon University, stated, "No significant relation was found between publication index and either teaching quality or student evaluation of teaching"(p. 228). Voeks, in an earlier study (1962) at the University of Washington, found the ratings faculty members received from their students bore no significant relationship to their publication rate or their membership in the University's Research Society. Stallings and Singhall (1969) also found no significant relationship between research productivity and student evaluation of courses at Indiana University and the University of Illinois.

Description of the Study

The empirical studies mentioned above, like the statements of opinion, have not provided a clear answer to the question. They have, however, suggested techniques upon which we built for this study. Dressel's study suggested the utility of using the American Council on Education's rating of graduate departments as an index of research and publication; Bresler, Hayes, and Voeks all examined teaching where it is operationalized: in the classroom; Hayes examined the rank of faculty and the level at which instruction was offered, but did not, along with Bresler and Voeks, examine how students' ratings differed by level of instruction. And these studies, while including different departments, did not differentiate among the areas of knowledge. This study thus builds upon these past studies, yet examines the

---

[4]Fiedler and Biglan (1969) in a study of departments at the University of Illinois found a correlation of .54 between the ACE rating and the size of departmental facultie suggesting ranking may be a function of size: larger departments are more visible. Th larger departments at the University of Washington, likewise, were in the more highly ranked groupings. See the staff profiles in Appendix E.

relationships from new vantage points. For an index of quality in research we used the most recent ACE study on the quality of graduate faculties; for an index of quality instruction we designed a new student rating instrument.

Quality of Graduate Faculties

Although there have been a number of national studies of departmental reputation over the past fifty years (Hughes, 1928: American Council on Education, 1934; and Keniston, 1959), the most comprehensive were undertaken in 1964 and 1969 by the American Council on Education (ACE): Cartter (1966), Rosse and Anderson (1970).

Raymond Hughes is credited with the first assessment of quality in graduate education. In 1924, while he was president of Miami University (Ohio), in order to more properly advise Miami undergraduates, requested his faculty to draw up a list of recognized scholars in each of twenty fields of study, to which he then sent questionnaires. This technique of assessment based upon the opinions of knowledgeable experts was repeated in 1957 by Hayward Keniston at the University of Pennsylvania. Keniston was dean of the graduate school and wished to assess his school's ranking with the twenty-five institutions belonging to the Association of American Universities. Both of these reports were used extensively for comparison with the recent ACE studies.

The most recent ACE report (Roose and Anderson, 1970) is used in this study, and shares with earlier studies the general limitation of being essentially a subjective estimate of quality. Quality is an elusive attribute, not easily subjected to objective measurement, thus the ACE sought to obtain the best judgments of quality by including a large sample of informed scholars at a substantial number of institutions.[5] The judgments of the respondents were in general agreement when compared with "objective" measures of quality such as size of library, publication record of faculty, level of faculty salaries, number of Guggenheim fellows, and so on; however, any objective measure taken as a single index was found to be misleading--for example, for average faculty compensation in 1965 Harvard was first, Parsons College second; Berkeley has little endowment; California Institute of Technology has a small library and Michigan and M.I.T. have no nobel laureates. Thus, when one measures quality there may be no more accurate operational definition than the opinion of informed observers.

Another limitation in using the ACE report as a measurement of departmental research reputation is the problem of time lag: (a) between change in a department and national awareness of this change;[6] and (b) between the date of the ACE study and the date of our student evaluation. Despite this limitation, the ACE report seemed an adequate enough measurement of current departmental reputation for the purposes of this study.

Using the ACE rankings as a basis, then, the departments at the University of Washington are ranked and grouped as presented in Tables 1 and 2. Table 1 illustrates a rank ordering and three groups of departments: "higher," "medium," and "lower" according to their rating by faculty in the national sample. The grouping was

---

[5] 4,008 faculty at 106 institutions in 1964, and 6,693 faculty at 130 institutions in 1969.

[6] Approximately half of the faculty appointments at the University were made within the last five years.

Table 1

Rank of University of Washington Departments Based Upon A.C.E. Ratings (1970)

Analytical Group	Departments	Percent Who Rated Distinguished and Strong	Percent Who Rated Good and Adequate
Higher	Geography	58	36
	German	49	28
	Mathematics	40	33
	English	39	28
Medium	Sociology	36	40
	Anthropology	29	58
	Physics	25	34
	Chemistry	22	38
	Psychology	17	45
	French	16	38
Lower Group	Economics, Geology, Mechanical Engineering, Music, Philosophy, Political Science, Spanish	*	*

*Alphabetical

Table 2

Rank of University Departments by Area of Knowledge*

Departmental Group	Area of Knowledge		
	Social Sciences	Humanities	Physical Sciences
HIGHER	Geography	English German	Mathematics
MEDIUM	Sociology Anthropology Psychology	French	Chemistry Physics
LOWER	Economics Political Science	Music Philosophy Spanish	Geology Mechanical Engineering

*Departments in the natural sciences were not included in the study for lack of departments which were equivalent in name to those ranked in the report.

arrived at by placing in the "higher" category the four departments whose graduate faculty was rated by the majority of respondents as distinguished and strong; the "medium" category consists of those departments rated by the majority of respondents as good and adequate; and the final category, which we have labeled "lower," is an alphabetical listing of departments whose average rating was between adequate and good.

Even though we have grouped certain departments as lower, this is clearly not an indication of low status among departments in a given field of knowledge. For example, Economics and Spanish, among graduate departments of economics and Spanish nationally, are among the top third of ranked departments when considered within their disciplines. The point is, all the departments included in this study are highly regarded when considering the 130 institutions included in the 1970 ACE report, but for the purposes of this study, among highly regarded departments we have selected for categories of analysis those which are higher, medium and lower, compared with each other.

In deciding on how to measure the quality of teaching, we examined in considerable detail the results from a number of recent studies, which we will not seek to review here, other than to point out we found that there seems to be substantial agreement among researchers, faculty, and students, that effective teaching includes qualities such as competence, caring, energy, imagination, enthusiasm, preparation, and so on. In fact, while the terms used from one study to another vary considerably, there nevertheless is a high degree of uniformity found in the central attributes of both good teacher and teaching. For example, each study or opinion we examined included a dimension of dynamism, enthusiasm, energy, spirit, zeal or whatever a particular researcher chose to call this attribute. Anyone wishing to examine the better publications on the evaluation of teaching effectiveness is referred to Eble (1970), Hildebrand, Wilson and Dienst (1971) and Miller (1972).

While we used existing publications to determine which concepts to include in our rating scale, we felt that the format and scope of items of the existing scales were inadequate for use in this study. (For our instrument, see Figure A.)

In format, we wanted items so designed that there would be maximum agreement as to the meaning of a response. On most instruments, the respondent is asked to rate instruction by selecting a number along a continuum of paired attributes, for example:

	Low						High
Enthusiasm	1	2	3	4	5	6	7

We felt it was important to label the numbers with verbal descriptions to insure that a given number meant the same degree of the attribute to all respondents. In doing this, we followed the descriptive technique of Russel Eidsmoe's instrument (Appendix A

In scope of items, we wanted as few items as possible, including some measuring the extent to which research is being incorporated into teaching (since this is one of the key arguments for saying that teaching and research go hand in hand) as well as some measuring generally accepted attributes of good teaching which are unrelated to research.

Of the eight items in our instrument, three were research-oriented: Knowledge of Subject (No. 2), Currentness of Material (No. 4) and Use of Own Research (No. 7). The remaining five items represented general qualities of teaching unrelated (or

SCALE FOR STUDENT ASSESSMENT OF TEACHING

The following scale consists of qualities which instructors may possess to varying degrees. Please rate the instructor of the course listed above on each of the qualities. Each quality is divided into three sections, and each section into three degrees, numbered accordingly from 1 to 9. In rating, circle the number which best describes your instructor.

	1 2 3	4 5 6	7 8 9
1 ENTHUSIASM FOR SUBJECT	Is definitely enthusiastic about subject matter of course.	Is at times enthusiastic about subject matter.	Seems to teach course without enthusiasm.
2 KNOWLEDGE OF SUBJECT	Definitely an expert; well-read and experienced.	In good command of subject matter.	Unsure of subject matter; not well-read.
3 TOLERANCE	Encourages and accepts widely differing points of view.	Accepts views somewhat differing from his own.	Discourages differing points of view; intolerant.
4 CURRENTNESS OF MATERIAL	Keeps course up to date; presents results of current research or scholarship; introduces still-debateable issues.	Occasionally mentions current or recent research in the field.	Course preparation could have been made years ago.
5 PRESENTATION	Stimulating and exciting.	Moderately interesting.	Puts you to sleep.
6 AVAILABILITY TO STUDENTS	Is often available and easily approached for formal or informal contact in addition to posted office hours.	Reasonably friendly; can be approached after class and during posted office hours.	Resists student contact, rarely available after class; rarely keeps, or has no regular office hours.
7 USE OF OWN RESEARCH	Consistently incorporates own research into course.	Occasionally incorporates own research into course.	Little or no evidence of involvement in research.
8 PREPARATION	Shows definite evidence of careful preparation.	Shows some preparation.	Not well prepared.

perhaps inversely related) to research: Enthusiasm for Subject (No. 1), Tolerance (No. 3), Presentation (No. 5), Availability to Students (No. 6), and Preparation (No. 8).

In addition, Item 5 (Presentation) was chosen because this attribute, with similar descriptors, was used in the Hayes (1971) study to measure quality of teaching. By including this item in our instrument, we could test whether the finding of a correlation between quality teaching and research varies according to the item used to measure the quality of teaching.

Measuring Quality in a Subjective Survey

While the explicitness of the scales developed for this study helps reduce unreliability due to variations in interpretation of meanings of scale terms and gradations, the question remains whether student respondents are qualified to judge their instructors.

Several studies mentioned by Costin et.al. (1971) have shown that students are reliable sources of judgments about faculty. In addition, if intellectual ability is any indication of ability to rate instructors, then our respondents are well qualified as raters, since the academic quality of the students at the University of Washington is quite high. In a recently reported comparison of selectivity among the larger universities, Astin (1970) reported that the University of Washington was second only to Harvard, and ranking above the Universities of Michigan and California, in the quality of its undergraduate student body. Thus, the respondents are among the most able in our universities.

Considering both the ability level of the respondent and the clarity of the rating instrument, it is our opinion that this technique yields as reliable an indicator as can be presently devised in attempting to measure the elusive attribute of quality instruction.

Respondents

The respondents were chosen from the University of Washington FTE student body of thirty-four thousand. From the individual programs of studies, copies of which were obtained from the University Registrar, students taking courses in the 17 departments studied were randomly selected. For control purposes, an eighteenth grouping was selected which included a random selection of students from all departments of the University including the departments under study.

Because the intent of the study was to assess the quality of instruction throughout a department, and not the performance of individual instructors, only the department and level of instruction were identified on the to-be-returned questionnaire. Since the study of Hildebrand, et.al. did not find the rating of instruction to be affected by the academic rank of the instructor, the number of courses previously taken in the same department, the class size, the student's major or whether or not the course was required, we did not seek these data on the courses evaluated or on

Table 3

Response Rates

Department	Number Mailed*	Number Returned	Percent Returned
Geography	70	39	55.7
German	70	44	57.1
Mathematics	65	33	52.3
English	70	33	47.1
Sociology	70	39	55.7
Anthropology	66	37	56.0
Physics	68	38	55.8
Chemistry	70	38	54.2
Psychology	70	38	54.2
French	62	30	48.4
Economics	68	41	58.8
Geology	61	23	45.1
Mechanical Engineering	56	39	69.6
Music	65	32	49.2
Philosophy	70	34	48.5
Political Science	61	25	40.9
Spanish	53	25	47.1
All University	90	55	51.1
	1106	643	58.1

* Differences in number mailed out reflect the size of the department. While 70 was the number aimed for in each department, our method of random selection made it impossible to select 70 students in the smaller departments.

the student respondents (Hildebrand, et.al., 1971). The cover letter indicated to the respondent the specific class selected for rating (see Appendix B). An addressed, business-reply envelope completed the mailing.

A total of 1106 questionnaires were mailed, and 643 of them (58.1 percent) were returned, forty of which could not be used for the following reasons: illegibility of marks (9), questionnaire inappropriate to format of class (11), and other reasons such as late return (20). The remaining 603 constitute the data sample.

Twenty students included extensive remarks in addition to their requested ratings. Seven critical remarks were based on the position that no objective instrument could ever measure the subjective art of teaching. Notes of approval generally expressed relief that "someone is finally doing something about university teaching." The most frequently written remark, found on twelve questionnaires, was that this method of data collection could not be applied to the seminar-style class, found most frequently at the graduate level of instruction. At least thirty respondents asked for copies of the research report when completed. No apparent problem was encountered in the mechanics of filling out the questionnaire. Questionnaires were mailed so as to arrive in the hands of respondents towards the end of the term, but at least two weeks before the pressure of final exams.

Table 3 illustrates the sample sizes, number of returns and response rates. Non-responses are not considered to be a serious problem in interpreting the results, since we are less interested in assessing the actual level of teaching performance than making comparisons among departments by their ranking for research. Thus, if there is a bias among non-respondents it is likely to bias the results evenly across departments.[7]

Analysis of Data

The analysis of the data consists of the following: (1) an examination of the internal consistency of the scales of the questionnaire, (2) an examination of the differences in the quality of instruction as perceived by students at the three levels of instruction, (3) an examination of the relationship between A.C.E. ranking of a department and the quality of instruction as perceived by our respondents, and (4) an examination of differences in ranking among the major fields of study.

The Scales: Internal Consistency

Tables 4 through 7 illustrate how the items are related (or not related) to each other in inter-correlational matrices. The inter-correlations indicate patterns of relationships which are the basis upon which items are combined for indices in later analyses.

Table 4 includes all complete responses from 602 cases. Figure 3 is a "cluster analysis" based upon the correlations in Table 4 where higher correlations ($r \geq .60$)

---

[7] The concern among survey researchers because of non-response seems unwarranted. Few investigations which have examined the non-respondent have found cause to feel the interpretation given to results from a partial sample would be changed significantl by a 100 percent response, e.g., Cope, 1968.

Table 4

Correlation Matrix of Scales: Full Number of Cases N = 602

Enthusiasm	1.00							
Knowledge	.56	1.00						
Tolerance	.35	.36	1.00					
Currentness	.49	.50	.40	1.00				
Presentation	.70	.57	.47	.53	1.00			
Availability	.38	.28	.44	.30	.36	1.00		
Research	.45	.50	.29	.54	.50	.32	1.00	
Preparation	.52	.62	.41	.44	.62	.34	.44	1.00
	Enthus-iasm	Know-ledge	Toler-ance	Current-ness	Presen-tation	Avail-ability	Research	Prepara-tion

Figure B

Cluster Analysis of Data from Table 4

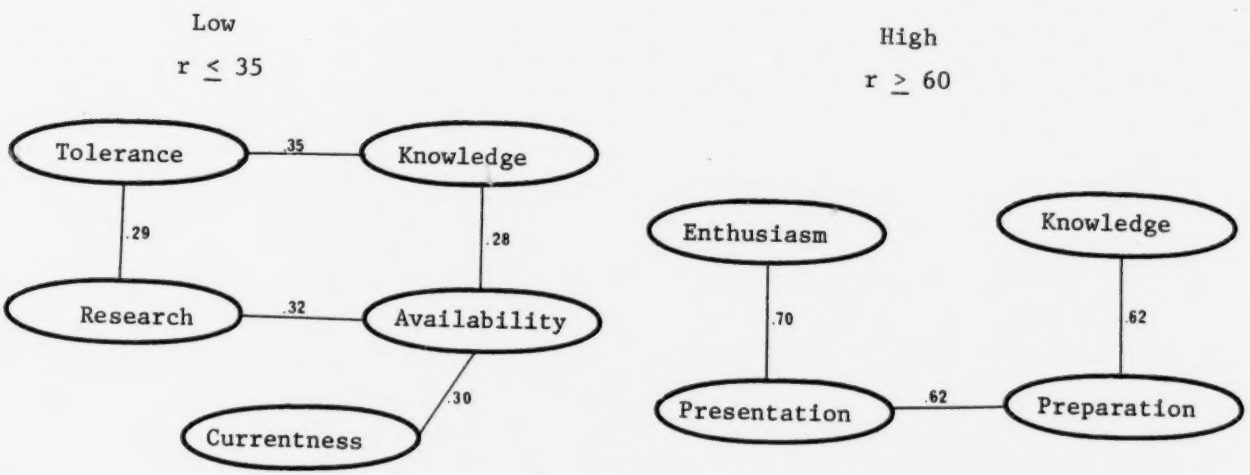


Table 5

Correlation Matrix of Scales: Lower Division 209

Enthusiasm	1.00							
Knowledge	.54	1.00						
Tolerance	.40	.37	1.00					
Currentness	.39	.40	.43	1.00				
Presentation	.68	.53	.47	.49	1.00			
Availability	.26	.22	.37	.27	.29	1.00		
Research	.39	.44	.29	.50	.45	.28	1.00	
Preparation	.49	.61	.41	.43	.58	.34	.39	1.00
	Enthus-iasm	Know-ledge	Toler-ance	Current-ness	Presen-tation	Avail-ability	Research	Prepara-tion

Figure C

Cluster Analysis of Data from Table 5

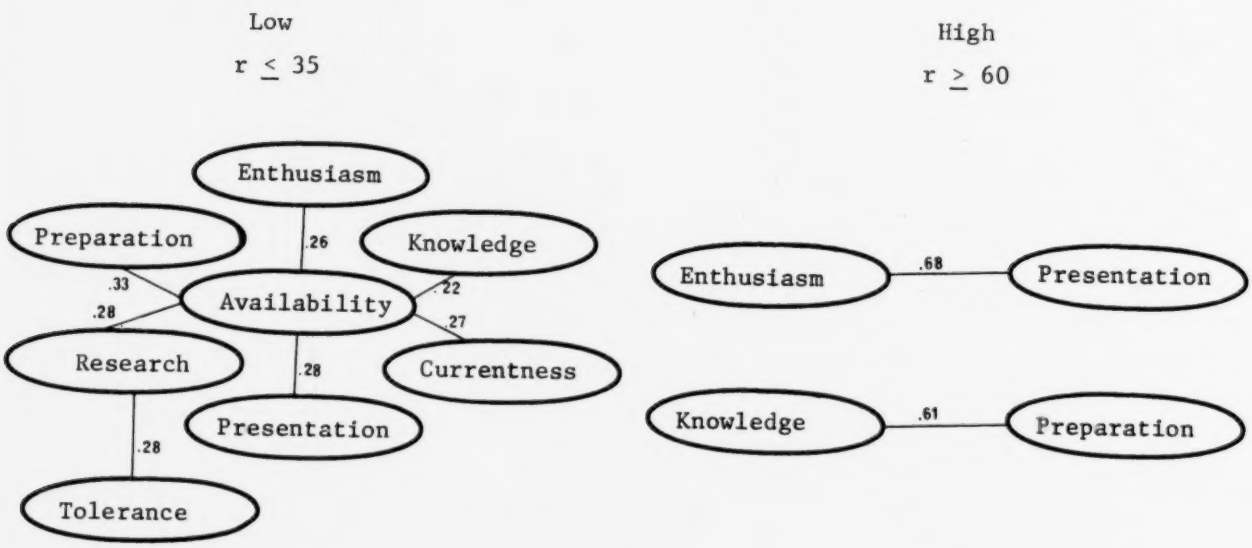


Table 6

Correlation Matrix of Scales: Upper Division N = 226

Enthusiasm	1.00							
Knowledge	.52	1.00						
Tolerance	.33	.32	1.00					
Currentness	.54	.48	.45	1.00				
Presentation	.69	.54	.49	.57	1.00			
Availability	.48	.32	.47	.35	.43	1.00		
Research	.44	.47	.26	.48	.51	.31	1.00	
Preparation	.51	.64	.37	.42	.62	.30	.46	1.00
	Enthus-iasm	Know-ledge	Toler-ance	Current-ness	Presen-tation	Avail-ability	Research	Prepara-tion

Figure D

Cluster Analysis of Data from Table 6

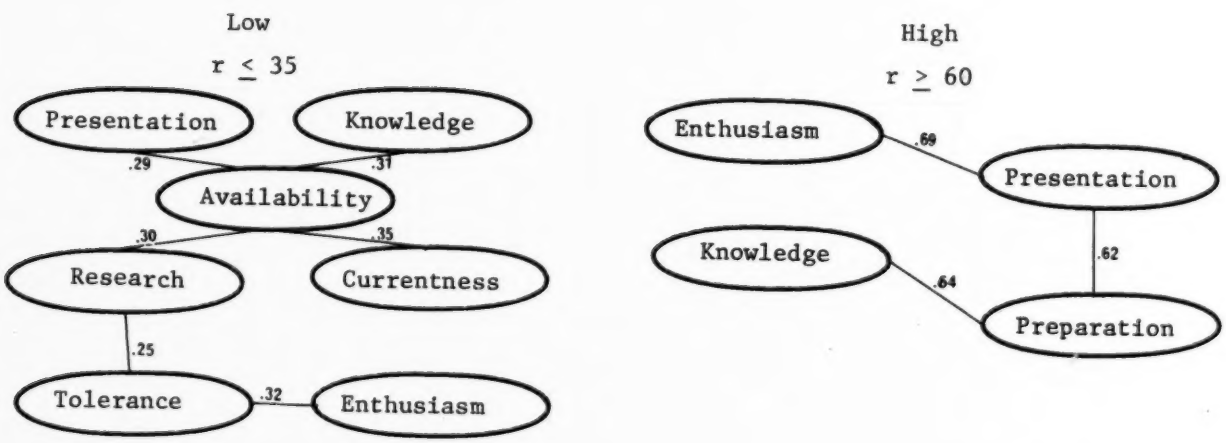


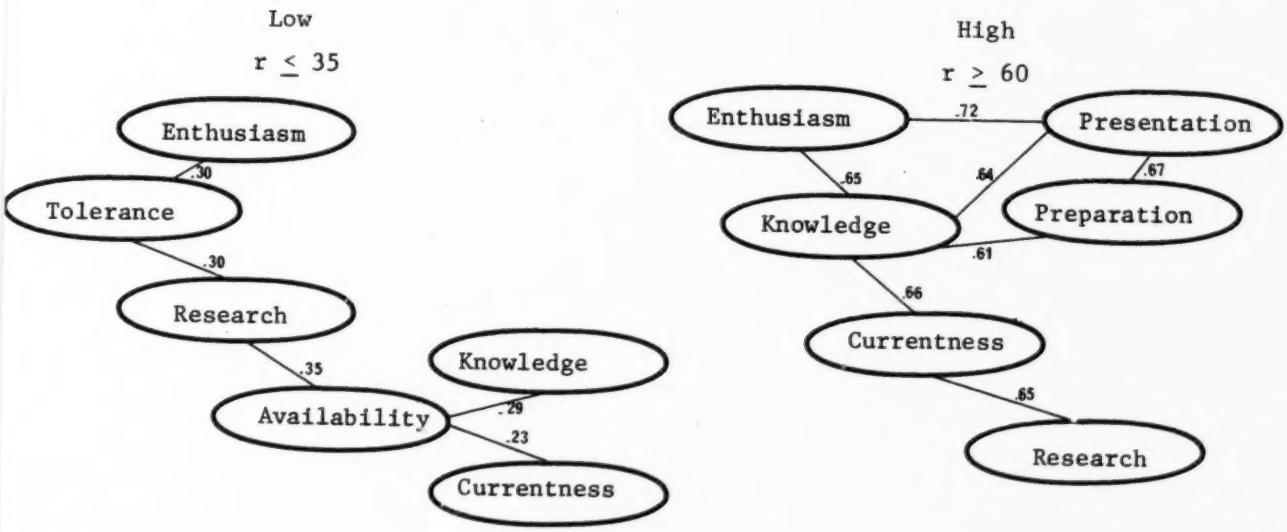
Table 7

Correlation Matrix of Scales: Graduate $N_r = 168$

Enthusiasm	1.00							
Knowledge	.65	1.00						
Tolerance	.31	.39	1.00					
Currentness	.56	.66	.26	1.00				
Presentation	.72	.64	.45	.53	1.00			
Availability	.41	.29	.50	.24	.36	1.00		
Research	.54	.57	.31	.65	.55	.35	1.00	
Preparation	.58	.61	.47	.50	.68	.41	.48	1.00
	Enthusiasm	Knowledge	Tolerance	Currentness	Presentation	Availability	Research	Preparation

Figure E

Cluster Analysis of Data from Table 7



and lower correlations ($r \leq .35$) are "clustered" for the purpose of illustration. As would be expected, scales such as Enthusiasm and Presentation are closely related, as are Knowledge and Preparation, while Presentation is related to both Enthusiasm and Preparation. Another way of stating these relationships is to say that respondent who found the presentation to be stimulating also believed the instruction was careful prepared and delivered with enthusiasm.

More interesting perhaps than the highly related scales are those without high correlations. The three research-oriented scales (Knowledge, Research and Currentness) are not found to be highly related to each other, but, more importantly, are virtually unrelated to the scales of Tolerance and Availability.

Correlations in Tables 5 and 6 are much like those in Table 4, illustrating again the relatedness of the scales for Enthusiasm, Presentation, Knowledge and Preparation for both levels of undergraduate instruction. However, in Table 7 the correlations based upon graduate responses indicate a new clustering of relationships. For the first time the three research oriented scales are found to have high inter-correlatio

Mean Scale Ratings by Level of Instruction

While data in Tables 4 through 7 illustrate how scales are interrelated at the three levels of instruction, Table 8 illustrates mean faculty ratings. Low means are indicative of student perception of better teaching.

On the whole the respondents have given the faculty favorable ratings; most ratings are at the level of 2 or 3, and, as one can see from the instrument, ratings through 3 were written to describe desirable attributes of teaching. These ratings are consistent with other research which has shown that students tend to rate instructors generously (see Hildebrand, et.al., 1971, p. 11).

Aside from the fact that respondents tended to give favorable ratings to all the forms of instruction measured by the eight scales, most favorable overall ratings were given to the instructors' enthusiasm and knowledge for and about the subject taught (means of 2.59 and 2.52). Least favorable overall ratings went to the attributes of a stimulating presentation of the subject matter ($\bar{X} = 3.81$) and the introduction of the instructors' own research ($\bar{X} = 3.48$).

The most significant portion of these data would appear to be the consistent and marked increased favorability of the rating given to instruction at the lower division to graduate levels of instruction as compared on the three research scales: Knowledge (#2), Currentness (#4), and Research (#7). None of the other five scales indicate an equal amount of improvement, and in only one of the other scales is the improvement as consistent (Availability) as it is in the three research scales. Figure F illustrates the changes according to levels of instruction for the eight scales. (For further analysis of these data, see Appendix F, Tables 1 and 2.)

Relationship Between the A.C.E Rank and Student Ratings

Table 9 shows the mean ratings on all eight scales given to instruction in the seventeen academic departments, which are ranked according to the A.C.E. ratings (except for those in the "lower" category, which are listed alphabetically). In

Table 8

Mean Scale Ratings for Instruction in
in Lower, Upper and Graduate Level Courses

Scale	Means			
	Lower Division	Upper Division	Graduate	Combined
Enthusiasm	2.65	2.78	2.34	2.59
Knowledge	2.71	2.67	2.19	2.52
Tolerance	3.28	3.28	2.91	2.95
Currentness	3.37	2.88	2.30	2.73
Presentation	3.90	4.00	3.65	3.81
Availability	3.20	3.02	2.75	2.86
Research	4.21	3.84	3.11	3.48
Preparation	2.81	2.92	2.77	2.81
	N = 209	N = 226	N = 168	N = 602

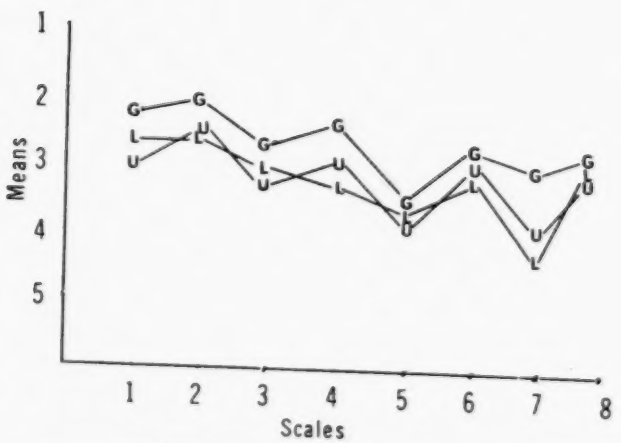


Fig. F. MEANS BY LEVEL OF INSTRUCTION

addition to the means for the scales there is a department average, calculated from the scale means. A range and number of respondents completed the table.

The first observation concerning these data, comparing the national (A.C.E) ranking with the teaching quality ranks ($\bar{X}_{1-8}$ and $\bar{X}_R$), is that there does not appear to be a clear relationship between a high national rating and the quality of instruction as perceived by students; in fact, two departments (French and Geology) with the highest overall ranking for teaching are in the lower half of the ranking for research. The rank-order correlations between A.C.E rankings and teaching quality rankings were .16 ($\bar{X}_{1-8}$) and -.07 ($\bar{X}_R$), both insignificant.

While the overall rankings for research and for instruction as illustrated in Table 9 do not present any consistent relationship for the seventeen departments, there is a slight suggestion that within the social sciences there may be a more consistent relationship than appears to be true of either the physical sciences or of the humanities. The student ratings and the ranks for the departments in the social sciences are illustrated in Table 10 and Figure G. One can see here a slight trend toward better student evaluations in the higher-ranked departments for questions 1, 2, 4, and 7. That questions 2, 4, and 7 were the research-oriented questions indicates a possibility that, for the social sciences, quality research may result in more effective teaching. (The rank-order correlation for social sciences between national rank and $\bar{X}_R$ ranking was +.59; however, the limited number of cases makes this statistically insignificant.)

Table 11 illustrates another ranking system which was employed to determine if research rankings might be related to the student evaluation rankings. The departmental ranking in this table is calculated from its standing in relation to departments in the <u>same</u> discipline; thus, Geography ranked first among departments included in this study at the University and second when its percentile was used to rank it in comparison to the 34 departments of Geography in the national sample; likewise, Physics, which is ranked seventh at the University for this study, is also ranked second when the basis for ranking becomes its percentile ranking as compared to departments of Physics. The purpose of this new ranking system was to adjust for any "generosity factor" that may have been operating in the former ranking system. In that system, a department's rank was based on the absolute score it received from other colleagues in the same discipline, so that a difference in the generosity of the members of two disciplines might account for the difference in scores received by the respective University of Washington departments. By re-ranking the departments according to their percentile rank in their discipline, we hoped to lessen this generosity factor.

Somewhat different correlations between national ranking and teaching quality rankings resulted when this new national ranking was used; rho became -.37 for $\bar{X}_{1-8}$ which is still statistically insignificant, and -.56 for $\bar{X}_R$ which is significant at the .05 level.

In order to try to interpret this correlation, it is helpful to look at the departments by field of study. Whereas with the previous ranking system, social sciences differed from the humanities and physical sciences in showing a more pronounced relationship between research and teaching, with the new ranking system, the field of study that stood out was the physical sciences (see Table 12). It seems likely that the overall negative correlation was strongly influenced by the rankings of the physical sciences, which received higher rankings in the re-ordering, yet received nearly the lowest student rankings for research (with the exception of Geology, for which the converse was true).

Table 9

Mean Student Ratings for Departments

	Departments	SCALES+ Enthusiasm	Knowledge	Tolerance	Currentness	Presentation	Availability	Research	Preparation	Range	$\bar{X}_{1-8}$	$\bar{X}_R$	RANK° X_{1-8}	RANK° X_R	N
HIGHER	Geography	2.30	2.23	3.21	2.46	3.87	2.73	2.84	2.89	2.2-3.8	2.82	2.51	5	3	39
	German	1.97	2.50	3.46	2.97	3.54	2.57	3.31	2.77	1.9-3.4	2.99	2.93	8	8	44
	Mathematics	2.72	3.06	3.38	3.72	3.68	3.62	4.93	3.15	2.7-4.9	3.53	3.90	15	16	33
	English	2.27	2.50	2.58	3.34	3.68	2.93	4.03	2.54	2.2-4.0	2.98	3.29	7	12	33
MEDIUM	Sociology	2.52	2.48	2.72	2.65	3.86	2.48	3.37	2.97	2.4-3.8	2.88	2.83	6	7	39
	Anthropology	2.37	2.52	2.70	2.27	3.33	3.72	2.91	2.11	2.1-3.7	2.74	2.56	3	4	37
	Physics	2.71	2.21	3.12	3.25	3.84	2.81	5.09	2.32	2.2-5.0	3.17	3.51	11	14	38
	Chemistry	2.76	2.29	3.81	3.32	4.21	3.22	4.41	3.02	2.2-4.4	3.38	3.34	14	13	38
	Psychology	2.58	2.81	3.48	2.40	3.54	3.55	3.59	2.75	2.4-3.5	3.09	2.94	10	9	38
	French	2.31	2.03	2.53	2.22	2.71	2.60	2.00	2.34	2.0-2.7	2.34	2.08	1	1	29
LOWER†	Economics	3.31	3.29	3.78	2.88	4.88	3.41	4.42	3.50	2.8-4.8	3.68	3.53	16	15	44
	Geology	2.39	1.86	2.65	1.59	3.39	2.33	2.85	2.00	1.5-3.3	2.38	2.10	2	2	23
	Mech. Eng.	3.20	3.07	3.10	4.30	4.63	2.82	5.05	3.53	2.8-5.0	3.71	4.14	17	17	39
	Music	2.34	2.03	4.03	3.34	4.31	3.70	3.65	3.12	2.0-4.3	3.32	3.01	13	10	32
	Philosophy	2.70	2.20	2.75	2.70	3.73	2.48	3.14	2.63	2.2-3.7	2.79	2.68	4	5	34
	Political Sci.	3.00	2.16	3.04	2.32	4.68	3.04	3.68	3.12	2.1-4.6	3.23	2.72	12	6	25
	Spanish	2.24	2.76	3.08	2.87	3.58	2.83	2.85	2.80	2.2-3.8	3.00	3.26	9	11	25
	UW	2.68	2.60	3.21	2.85	3.49	2.76	3.23	2.57	2.5-3.4	2.92	2.89	NA	NA	55
	Scale No.	1	2_R	3	4_R	5	6	7_R	8						

Table 9

(con't.)

+ Scales have a range from 1 to 9, with 1 best.

† Alphabetical order

° While ranks are used in assessing the student ratings as they relate to the research rank, it is hoped that the reader will not attach too much importance to a specific rank; the rank for teaching should be regarded in terms of a quality range rather than a specific point to which one might infer a pecking order. In fact a tenth of a point in either direction for many departments would result in a change of several ranks in the ordering.

Table 10

Mean Student Ratings for Social Science Departments

	Scales										Rank	
	1	2	3	4	5	6	7	8	$\bar{X}_{1-8}$	$\bar{X}_R$	$\bar{X}_{1-8}$	$\bar{X}_R$
Geography	2.30	2.23	3.21	2.46	3.87	2.73	2.84	2.89	2.82	2.51	2	1
Sociology	2.52	2.48	2.72	2.65	3.86	2.48	3.37	2.97	2.88	2.83	3	4
Anthropology	2.37	2.52	2.70	2.27	3.33	3.72	2.91	2.11	2.74	2.56	1	2
Psychology	2.58	2.81	3.48	2.40	3.54	3.55	3.59	2.75	3.09	2.94	4	5
Economics	3.31	3.29	3.78	2.88	4.88	3.41	4.42	3.50	3.68	3.53	6	6
Political Science	3.00	2.16	3.04	2.32	4.68	3.04	3.68	3.12	3.23	2.72	5	3

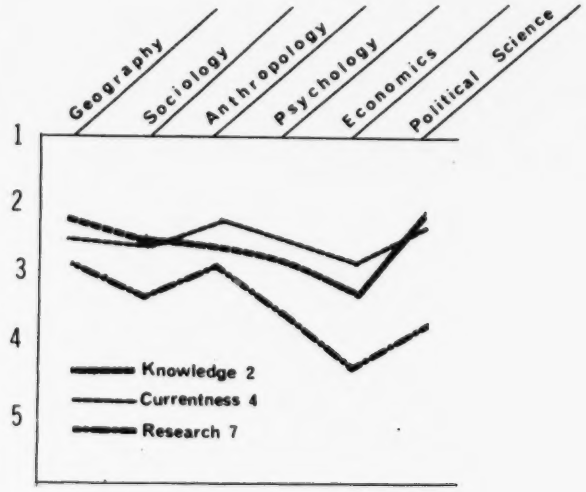


Fig G. GRAPH OF SCALES 2, 4, and 7.

Table 11

Rank Order of Departments by A.C.E. Rank and Department Means

	All Disciplines A.C.E.	Own Field A.C.E.			Ranked on Student Ratings	
		$x/\Sigma x$	%ile	R	$\bar{X}_{1-8}$	$\bar{X}_{R}$
Geography	1	6/34	82	2	5	3
German	2	12/48	75	8	8	8
Mathematics	3	18/102	82	2	15	16
English	4	12/92	87	1	7	12
Sociology	5	13/73	82	2	6	7
Anthropology	6	16/42	62	16	3	4
Physics	7	20/113	82	2	11	14
Chemistry	8	24/125	81	6	14	13
Psychology	9	27/110	75	8	10	9
French	10	19/63	71	14	1	1
Economics	11	19/91	80	7	16	15
Geology	11	22/69	69	15	2	2
Mechanical Engineering	11	20/71	73	10	17	17
Music	11	18/43	59	17	13	10
Philosophy	11	19/65	72	13	4	5
Political Science	11	21/74	73	10	12	6
Spanish	11	18/65	73	10	9	11

Table 12

Rank Order of Departments by Areas of Knowledge
by ACE Rankings and Mean Student Ratings

	ACE_1	ACE_2	Student Ratings $\bar{X}_R$
Social Science			
Geography	1	2	3
Sociology	5	2	7
Anthropology	6	16	4
Psychology	9	8	9
Economics	11	7	15
Political Science	11	10	6
Humanities			
German	2	8	8
English	4	1	12
French	10	14	1
Music	11	17	10
Philosophy	11	13	5
Spanish	11	10	11
Physical Science			
Math	3	2	16
Physics	7	2	14
Chemistry	8	6	13
Geology	11	15	2
Mechanical Engineering	11	10	17

Summary of Results

1. Respondents gave favorable ratings on all eight aspects of instruction measured in this study. The instructor's enthusiasm for and knowledgeability about his subject received the highest ratings. Rated lowest were the stimulating presentation and the introduction of the instructor's own research into the process of teaching.

2. Knowledge, Currentness, and Research, that is, the three research-oriented scales were rated consistently better by students taking advanced courses.

3. When departments were ranked according to their A.C.E. rating, no overall correlation between quality of research and quality of instruction was found; however, there did appear to be a slight positive relationship for the social sciences.

   When departments were ranked according to their rating in relation to other departments in their field, a slight overall negative correlation was found, which seemed to be most influenced by the strong negative relationships of the physical sciences.

4. More obvious than any overall correlations were the differences found among fields of study and the uniqueness of departments within these fields.

Implications

Because we built upon the results of previous research as we began this study we were initially inclined to believe our results would be somewhat definitive. However, as we have come to a fuller realization of the limits of our research design and a fuller understanding of the complexities of academic disciplines, we are more certain that what we have accomplished is another necessary preliminary examination of a very complex relationship.

We are satisfied that the survey instrument (the scales) is a useful means for an examination of teaching quality, especially the portion represented by the scales which were designed to examine the research-related aspects of instruction. We feel the instrument would be appropriate for use in future studies of university-level instruction. We were also satisfied with using the department as the unit of analysis, and with the Roose and Anderson report as a reasonable way to approximate the eminence of a department relative to departments in other institutions.

We were not satisfied, however, with our design of comparing departments within a university. Because of departmental differences and differences among fields of study, we feel a more definitive approach might be to examine the same department at different universities.

Our samples from each department are also quite small, ranging from as few as 23 respondents in Geology to 44 respondents in German, whereas the student course registrations in these departments approximated 1100 and 1250, respectively. Thus, we feel the mean ratings should only be interpreted as an estimate, and the resulting

rank ordering even more of an approximation, useful as a means of identifying tentative relationships, which would then be subject to more careful scrutiny.

The reader is also cautioned to recognize that even if negative findings were found in studies of university faculty with the dual roles the conclusion should not be drawn that the quality of instruction is better if the faculty does not have a commitment to research. Studies which included teaching quality comparisons between full-time university faculty with dual roles and full-time faculty with teaching roles (perhaps in other levels of the higher educational system) would provide information which is closer to testing the basic question: Is instruction benefited by having faculty fill dual roles? Then there is a related question: Is research benefited by scholars who must teach?

And finally, while the research was carried out at an institution which is similar to other large institutions, we should not be too hasty to generalize to other institutions which appear to be similar; substantial differences remain between institutions. Thus, some of the major state universities (Michigan, Minnesota, and Illinois, for example) appear to pay a good deal more attention to undergraduate teaching than other eminent state universities (such as Berkeley).

References

Anderson, Kay J. "The Ambivalent Department," Educational Record, Spring 1968, pp. 206-213.

Astin, Alexander W. "How Colleges are Rated," Change, November-December 1970, pp. 11 - passim.

Bresler, Jack B. "Teaching Effectiveness and Government Awards," Science, 160:3824, April 12, 1968, pp. 164-167.

Biglan, Anthony. Relationship Between the Characteristics of Academic Tasks and the Structure and Output of University Departments. University of Washington, Department of Psychology, Organizational Research, Technical Report 71-24, July 1971.

Cartter, Allan M. An Assessment of Quality in Graduate Education: A Comparative Study of Graduate Departments in 29 Academic Disciplines. Washington: American Council on Education, 1966.

Cartwright, Philip W. "The Economics of Deaning: The Care and Feeding of Homo Academicus," Western Economic Journal, 3:2, Spring, 1965; reprinted in A Sketch of University Decision-Making by the Associated Students of the University of Washington, January 1970.

Cope, Robert G. "Nonresponse in Survey Research as a Function of Psychological Characteristics and Time of Response," Journal of Experimental Education, 36:3, Spring, 1968, pp. 32-35.

Desmond, Richard L. "Faculty and Student Frustration Shaping the Future of the University," American Association of University Professors Bulletin, Spring, 1969, pp. 23-26.

Dressel, Paul L., F. Craig Johnson and Philip M. Marcus. The Confidence Crisis. San Francisco: Jossey-Bass, Inc., 1970.

Dwyer, Francis M. "A Review of Characteristics of Selected Criteria for Evaluating Teacher Effectiveness." Pennsylvania State University, Division of Instructional Services, January 1968.

Eble, Kenneth E. The Recognition and Evaluation of Teaching. Salt Lake City: Project to Improve College Teaching, 1970.

Fiedler, F. E., and Anthony Biglan. The First Job After Graduation as a Measure of Departmental Performance. Technical Report 74, Urbana, Ill.: Group Effectiveness Research Laboratory, April 1969.

Gross, Neal. "Organizational Lag in American Universities," Harvard Educational Review, 1963, 33, pp. 58-78.

Hayes, John R. "Research, Teaching, and Faculty Fate," Science, 172:3980, April 16, 1971, pp. 227-230.

Hildebrand, Milton, Robert Wilson, and Evelyn Dienst. Evaluating University Teaching. Berkeley: University of California Center for Research and Development in Higher Education, 1971.

Hughes, Raymond. "National Ranking of Graduate Schools," presented at the annual meeting of the Association of American Colleges 1925, and published in American Universities and Colleges, Washington, D.C.: American Council on Education, 1928.

Jencks, Christopher and David Riesman. The Academic Revolution. New York: Doubleday, 1968.

Katz, Solomon. "Administrative Commentary," Daily, Student newspaper at the University of Washington, April 4, 1970, p. 4.

Keniston, Hayward. Graduate Study in the Arts and Sciences at the University of Pennsylvania. Philadelphia: University of Pennsylvania Press, 1959.

Langen, Thomas. "Student Assessment of Teaching Effectiveness," Improving College and University Teaching, 14, 1966, pp. 22-25.

McKeachie, W. J., Lin Yi-Guang, and William Mann. "Student Ratings of Teacher Effectiveness," American Educational Research Journal, May 1971, pp. 435-445.

Miller, Richard I. Evaluating Faculty Performance. San Francisco: Jossey-Bass, Inc., 1972.

Roose, K. D. and C. J. Anderson. A Rating of Graduate Programs. Washington, D. C.: American Council on Education, 1970.

Rossi, Peter H. "Research, Scholars, and Policy Makers: The Politics of Large Scale Research," Daedalus, 93, 1964, pp. 1142-1161.

Sample, Steven B. "Inherent Conflict Between Research and Education," Educational Record, 33:1, 1972, pp. 17-22.

Sanford, Nevitt. Where Colleges Fail. San Francisco: Jossey-Bass, Inc., 1967.

Smith, R. and Fred E. Fiedler. "The Measurement of Scholarly Work: A Critical Review of the Literature," Educational Record, Summer, 1971, pp. 225-232.

Stallings, William M. and Sushila Singhal. "Some Observations on the Relationships Between Research Productivity and Student Evaluations of Courses and Teaching," The American Sociologist, 5:2, May 1970, pp. 141-143.

Voeks, Virginia. "Publications and Teaching Effectiveness: A Search for Some Relationship," Journal of Higher Education, 33, 1962, p. 212.

APPENDICES

Appendix

A	Sample Page from Roose and Anderson, 1970
B	Eidsmoe Scale
C	Cover Letter
D	Scale for Student Assessment of Teaching
E	Staff Deployment by Department
F	Details on Scale Means

SOCIOLOGY — Social Sciences

Leading Institutions, by Rated Quality of Graduate Faculty

Rankings 1957	Rankings 1964	Rankings 1969	Institution	Percentage[a] of Raters Who Indicate: "Quality of Graduate Faculty" Is: Distinguished and strong	Good and adequate	All other	Insufficient information
			Twenty-one institutions with scores in the 3.0—5.0 range, in rank order				
6	1	1*	California, Berkeley	91	3	–	6
1	2	1*	Harvard	91	5	–	4
3	4	3	Chicago	86	6	–	7
2	3	4*	Columbia	83	11	–	5
4	5	4*	Michigan	84	8	–	8
12	6	6	Wisconsin	80	10	–	9
8	9*	7	North Carolina	67	24	–	9
11	11*	8	California, Los Angeles	51	33	–	16
5	7*	9*	Cornell	51	39	–	9
‡	18*	9*	Johns Hopkins †	47	31	–	21
13	13*	9*	Northwestern	46	37	1	15
‡	7*	9*	Princeton	48	34	–	17
9	13*	13*	Washington (Seattle)	36	40	–	24
10	15*	13*	Yale	43	45	1	11
7	9*	15*	Minnesota	35	50	1	14
‡	11*	15*	Stanford	37	49	–	14
‡	17	17*	Michigan State	32	54	–	13
‡	18*	17*	Texas †	23	48	1	27
‡	18*	19	Indiana †	23	56	1	20
‡	‡	20*	*Brandeis*	24	54	2	20
15	‡	20*	Pennsylvania †	20	51	1	28

Nine institutions with scores in the 2.5—2.9 range, in alphabetical order

Brown †
Duke †
Illinois
M.I.T.
N.Y.U.
Oregon
Southern California †
Vanderbilt †
Washington (St. Louis)

Sixteen institutions with scores in the 2.0—2.4 range, in alphabetical order

Buffalo †
Case Western Reserve † [b]
Colorado †
Florida State †
Iowa (Iowa City)
Massachusetts
Missouri †
New School
Notre Dame †
Ohio State
Penn State
Pittsburgh
Purdue
Syracuse †
Tulane †
Washington State

Italicized institutions were not included in the 1964 survey of this discipline.
* Score and rank are shared with another institution.
† Institution's 1969 score is in a higher range than its 1964 score.
‡ Not ranked.
a. Percentages add across; the sum may not total 100 because of rounding.
b. Institution rated in 1964 was Western Reserve University.

Sample Page from A Rating of Graduate Programs by Kenneth D. Roose and Charles J. Anderson, American Council on Education, 1970. Reproduced with permission.

A STUDENT'S RATING SCALE OF AN INSTRUCTOR

Instructor's name (Please print) _____ Course _____ Date _____

Each of the qualities listed below is divided into three sections. Each section is divided into three degrees and numbered accordingly from 1 to 9, 1 being the highest degree and 9 the lowest. In rating, draw a circle around the number which best describes your instructor.

Your fair and honest opinion is what really counts. Your instructor desires this rating for his own self-improvement.

ORGANIZATION OF COURSE	1 2 3 Well organized; shows thoughtful planning.	4 5 6 Some organization but not always clear.	7 8 9 Lacks organization; planning seems vague.
PREPARATION FOR EACH CLASS	1 2 3 Shows definite evidence of careful preparation.	4 5 6 Shows some preparation; average knowledge of course.	7 8 9 Not well prepared; knowledge inaccurate at times.
TEACHING SKILL	1 2 3 Produces steady interest in subject; creates real desire; keeps things moving.	4 5 6 Teaching procedure seldom changes; student interest moderate.	7 8 9 Classes tend to be dry and uninteresting; class period drags.
ENTHUSIASM AND INTEREST IN COURSE	1 2 3 Keeps up steady interest and enthusiasm. Inspires interest in subject.	4 5 6 Appears to be reasonably interested.	7 8 9 Seems to teach course without enthusiasm.
ASSIGNMENTS	1 2 3 Students understand the tasks of each new assignment. Students know what is desired.	4 5 6 Sometimes rather indefinite; without clear planning.	7 8 9 Usually hurriedly given; rather vague; sometimes very unreasonable.
JUDGMENT OF VALUES	1 2 3 Usually selects important ideals; broadens student view points.	4 5 6 Sometimes overlooks important points, spending time on insignificant details.	7 8 9 Frequently misses important ideas; overemphasizes trivial details.
CLASS DISCUSSION AND QUESTIONS	1 2 3 Questions challenging; demand sound thinking; discussions interesting and stimulating.	4 5 6 Questions rather easy and simple; memorized facts emphasized.	7 8 9 Discussion sometimes without purpose; discussions frequently ramble.
POISE AND SELF-CONFIDENCE	1 2 3 Well poised; sure of himself; not easily upset.	4 5 6 Seems embarrassed at times; fairly self-confident.	7 8 9 Easily upset; uncertain as to procedure; lacks confidence.
EXAMINATIONS	1 2 3 Questions thought-provoking; carefully selected; clear.	4 5 6 Questions usually factual; require little thinking.	7 8 9 Examinations poorly planned and managed.
SCHOLARSHIP	1 2 3 Excellent mastery of subject; has broad interests.	4 5 6 Knowledge fair but without depth.	7 8 9 Knowledge frequently inadequate. Instructor seems vague.
ABILITY TO CREATE STUDENT INTEREST	1 2 3 Usually keeps steady interest in subject; stimulates thinking.	4 5 6 Students have average amount of interest.	7 8 9 Classes drag and students are indifferent.
CLASSROOM MANAGEMENT AND DISCIPLINE	1 2 3 Efficient management; students orderly and attentive.	4 5 6 Satisfactory organization; few disciplinary problems.	7 8 9 Poor organization; many disciplinary problems.
SPEECH	1 2 3 Voice pleasant; speaks distinctly, fluently.	4 5 6 Speaks reasonably well.	7 8 9 Enunciation poor; makes frequent errors in speech.
TOLERANCE	1 2 3 Encourages students to express opinions even though they differ with the instructor's ideas.	4 5 6 At times appears to be disturbed and impatient when students oppose instructor's views.	7 8 9 Resents opposition; intolerant.
SENSE OF HUMOR	1 2 3 Possesses keen sense of humor.	4 5 6 Moderately humorous at times.	7 8 9 Shows little or no sense of humor; quite sober and serious.
PERSONAL APPEARANCE	1 2 3 Neatly and appropriately dressed; well groomed.	7 8 9 Appearance fair; makes average impression.	7 8 9 Careless in dress; untidy.
RELATIONSHIP BETWEEN STUDENTS AND INSTRUCTOR	1 2 3 Attitude of friendliness; feeling of mutual interest; easily approached.	4 5 6 Neither ill-will nor friendliness prevails; attitude somewhat indifferent.	7 8 9 Considerable spirit of antagonism between students and instructor.

PRINT your criticisms of the course. These will be very helpful for your instructor's self-improvement. Do not sign your name.

On the back of this form PRINT any annoying mannerisms your instructor has developed which should be corrected.

Copyright, 1970, by Dr. Russell M. Eidsmoe, Morningside College, Sioux City, Iowa 51106

Reproduced with permission.

Student Assessment of Teaching

FACULTY–STUDENT
RESEARCH TEAM:

Robert Cope
Judy Richardson
John Mc Millin

University of
Washington
M–217
Miller Hall

Seattle
Washington
98195

Phone
543–1891

May 19, 1972

Dear Student:

People involved in Higher Education are currently attempting to determine the characteristics of effective teaching. The person being taught is probably the best source of useful data; therefore, your help is needed in a study of some factors that make up effective teaching. The research is being conducted by a team of students and faculty here at the University of Washington.

According to enrollment records, you are presently a student in

and we are asking you to complete the enclosed brief questionnaire concerning the instruction in that course.

As you will see, the numerical code on the questionnaire indicates only the department and course level, thus assuring anonymity for both you and the instructor. Because of this anonymity, participation in the project could in no way effect your grade or reflect on the instructor, and is in no way a requirement for the class.

Please return the questionnaire in the enclosed postage-paid envelope as soon as possible. It is important for the research that we receive your response to each item, but you may leave blank an item you find inappropriate for any reason. Feel free to call or visit any of us if you have a question, or if you wish to have a report of the findings.

Thank you!

BR

df

SCALE FOR STUDENT ASSESSMENT OF TEACHING

The following scale consists of qualities which instructors may possess to varying degrees. Please rate the instructor of the course listed above on each of the qualities. Each quality is divided into three sections, and each section into three degrees, numbered accordingly from 1 to 9. In rating, circle the number which best describes your instructor.

	1 2 3	4 5 6	7 8 9
ENTHUSIASM FOR SUBJECT	Is definitely enthusiastic about subject matter of course.	Is at times enthusiastic about subject matter.	Seems to teach course without enthusiasm.
KNOWLEDGE OF SUBJECT	Definitely an expert; well-read and experienced.	In good command of subject matter.	Unsure of subject matter; not well-read.
TOLERANCE	Encourages and accepts widely differing points of view.	Accepts views somewhat differing from his own.	Discourages differing points of view; intolerant.
CURRENTNESS OF MATERIAL	Keeps course up to date; presents results of current research or scholarship; introduces still-debateable issues.	Occasionally mentions current or recent research in the field.	Course preparation could have been made years ago.
PRESENTATION	Stimulating and exciting.	Moderately interesting.	Puts you to sleep.
AVAILABILITY TO STUDENTS	Is often available and easily approached for formal or informal contact in addition to posted office hours.	Reasonably friendly; can be approached after class and during posted office hours.	Resists student contact, rarely available after class; rarely keeps, or has no regular office hours.
USE OF OWN RESEARCH	Consistently incorporates own research into course.	Occasionally incorporates own research into course.	Little or no evidence of involvement in research.
PREPARATION	Shows definite evidence of careful preparation.	Shows some preparation.	Not well prepared.

Staff Deployment by Department

	Ranked Faculty	Sub-Faculty	Research Assistants	Sponsored Research	S:F Ratio	F.T.E. Student			% of Faculty Time Spent on Research			
						L.Div.	U.Div.	Grad.	L.Div.	U.Div.	Grad.	Total
Higher Group												
Geography	14.00	4.00	3	33,762	28.5	264	96	32	5.2	14.8	17.2	8.2
German	18.00	6.36	1	2,922	19.7	203	93	47	3.0	16.5	21.7	8.6
Mathematics	64.13	12.79	1	315,119	25.2	1,266	239	109	17.6	28.9	29.1	19.8
English	71.52	15.75	-	5,510	29.3	1,085	567	220	7.2	12.2	11.7	9.2
Total	167.65	38.90	5	357,313	26.6	2,818	995	408	11.4	16.9	17.9	13.2
Middle Group												
Sociology	23.17	12.00	-	69,563	41.1	674	401	64	3.7	29.4	27.0	13.8
Anthropology	18.51	4.25	2	166,281	44.5	499	241	53	11.8	16.0	15.0	13.2
Physics	36.74	6.61	82	2,013,522	22.1	483	53	116	16.7	18.7	23.2	17.7
Chemistry	28.65	18.52	19	541,911	28.8	808	111	115	12.9	10.0	19.3	13.1
Psychology	37.03	15.13	20	429,676	36.3	824	441	152	11.9	25.4	27.6	17.4
French	11.75	4.42	1	5,840	17.4	145	54	21	3.7	12.0	15.9	6.5
Total	155.85	60.93	124	3,226,793	32.0	3,433	1,301	521	10.9	22.7	22.9	14.7
Lower Group												
Economics	27.45	6.00	4	44,632	28.1	454	228	55	16.9	22.5	27.5	19.2
Geology	17.17	4.26	4	308,410	20.2	215	60	42	12.0	13.9	16.7	12.8
Mech. Engr.	39.55	.99	6	169,450	15.8	144	208	64	7.5	7.7	8.1	7.7
Music	53.00	6.26	-	5,494	11.3	221	185	62	6.7	7.7	8.3	7.3
Philosophy	16.83	5.25	-	-	37.3	586	128	19	14.6	19.1	22.5	15.6
Pol. Science	24.44	5.65	5	74,206	38.4	337	443	68	22.5	23.9	35.5	23.5
Spanish	11.75	4.42	1	5,840	17.4	145	54	21	3.7	12.0	15.9	6.5
Total	190.19	32.83	20	608,032	22.2	2,102	1,306	331	14.0	17.4	19.4	15.5
All Groups	513.69	132.66	149	4,192,138	26.9	8,353	3,602	1,260	11.8	19.2	20.4	14.4
Averages												
Higher Group	41.91	9.73	1.25	89,328	25.7	705	249	102	8.3	18.1	19.9	11.5
Middle Group	25.98	10.16	20.67	537,798	31.7	572	217	87	10.1	18.6	21.3	13.6
Lower Group	27.17	4.69	2.86	86,861	24.1	300	187	47	12.0	15.3	19.2	13.2
All Groups	30.22	7.80	8.76	246,596	27.1	491	212	74	10.4	17.1	20.1	12.9

Notes on Appendix E

(1) Ranked faculty are shown in terms of Autumn Quarter 1971 faculty count, where each full-time faculty member = 1.00 faculty count. Ranked faculty includes: Professor, Associate Professor, Instructor, and Lecturer.

(2) Sub-faculty shown in terms of Autumn Quarter 1971 faculty count, where Part-time Instructors, Part-time Lecturers, and Pre-doctoral Teaching Associate II's are counted at the same rate as ranked faculty and Pre-doctoral Teaching Associate I's and Teaching Assistants are counted at half that rate (e.g., each half-time T.A. = .25 faculty count).

(3) Research Assistants are shown in terms of F.Y. 1971 head count and include Pre-doctoral Research Associates.

(4) Sponsored Research are based on F.Y. 1971 data and exclude teaching grants and fellowships.

(5) The student faculty ratio include both ranked and sub-faculty as a base and are based on a weighted F.T.E. student distribution. This weighting reduces all students to lower division equivalents where: lower division = 1.00; upper division = 1.33; graduate students = 2.86; and graduate students doing independent study, thesis, or dissertations = 4.00.

(6) F.T.E. students were developed from Autumn Quarter 1971 student credit hours (SCH). They were computed on the basis that 15 undergraduate SCH's per undergraduate student and 10 graduate SCH's per graduate student.

(7) The percent of faculty time spent on research was computed from Autumn Quarter data and developed from a comparison of the percent of time spend on research by faculty rank and the teaching breakdown by rank.

(8) Data for the Departments of French and Spanish were calculated on the assumption that each represented approximately one-third of the Department of Romance Language and Literature.

Tables 1 and 2 (Appendix F) help answer the question, "Is the quality of instruction offered by levels of instruction related to the ranking of the departments for research?" Again the answer appears to be "no." There is no clearly consistent relationship for the scales in either table, except as noted earlier for the improved ratings on research oriented scales as the level of instruction increases. There is, however, a hint of a curvilinear relationship with the most favorable ratings on the eight scales received by departments in the medium ranking (note the asterisks on Table 1); fifteen out of a possible twenty-four best ratings were received by departments in the medium group.

The data in Table 2 are particularly important in regard to a central question of this study, "Does a national reputation among peers for quality of faculty result in improved instruction as perceived by students, particularly as instruction is related to the introduction of recent discoveries, theories, interpretations, i.e., research?" Again, these data do not indicate any relationship.

Table 1

Means on Scales by Grouped Rank of Departments and Level of Instruction

Departments	Scales	Lower Division	Upper Division	Graduate
H	1	2.15*	2.43*	2.30
I	2	2.71	2.60	2.30
G	3	3.30	2.86*	3.38
H	4	3.25	3.31	2.59
E	5	3.88	3.59*	3.58
R	6	2.82*	3.08	2.88
	7	3.74*	3.84	3.44
	8	3.01	2.92	2.51*
		(53)	(53)	(43)
M	1	2.48	2.82	2.23*
E	2	2.47*	2.54*	2.09*
D	3	2.89*	3.46	2.72*
I	4	3.14*	2.83	1.86*
U	5	3.30*	4.03	3.46*
M	6	3.37	2.96*	2.84
	7	4.02	3.83*	2.68*
	8	2.25*	2.90*	2.63
		(80)	(85)	(54)
L	1	3.06	2.86	2.42
O	2	2.80	2.69	2.12
W	3	3.45	3.32	2.95
E	4	3.68	2.73*	2.37
R	5	4.40	4.39	3.88
	6	3.23	3.11	2.55*
	7	4.57	3.93	3.20
	8	3.16	2.93	3.04
		(76)	(83)	(63)

* Most favorable scale rating based on column comparisons.

Table 2

Means on Research Scales by Grouped Rank of Departments and Level of Instruction

Departments	Scales	Lower Division	Upper Division	Graduate
Higher	2	2.71	2.60	2.30°
	4	3.25	3.31°	2.59°
	7	3.74*	3.84	3.44°
		(53)	(53)	(43)
Medium	2	2.47*	2.54*	2.09*
	4	3.14*	2.83	1.86*
	7	4.02	3.83*	2.68*
		(80)	(85)	(54)
Lower	2	2.80°	2.69°	2.12
	4	3.68°	2.73*	2.37
	7	4.57°	3.93°	3.20
		(76)	(83)	(63)

* Most favorable scale rating based upon column comparisons.
° Least favorable scale rating based upon column comparisons.